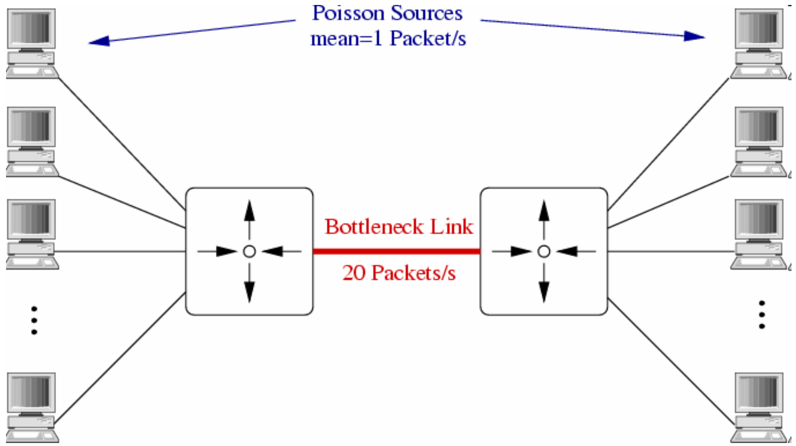
Modelos de Colas Poissonianos

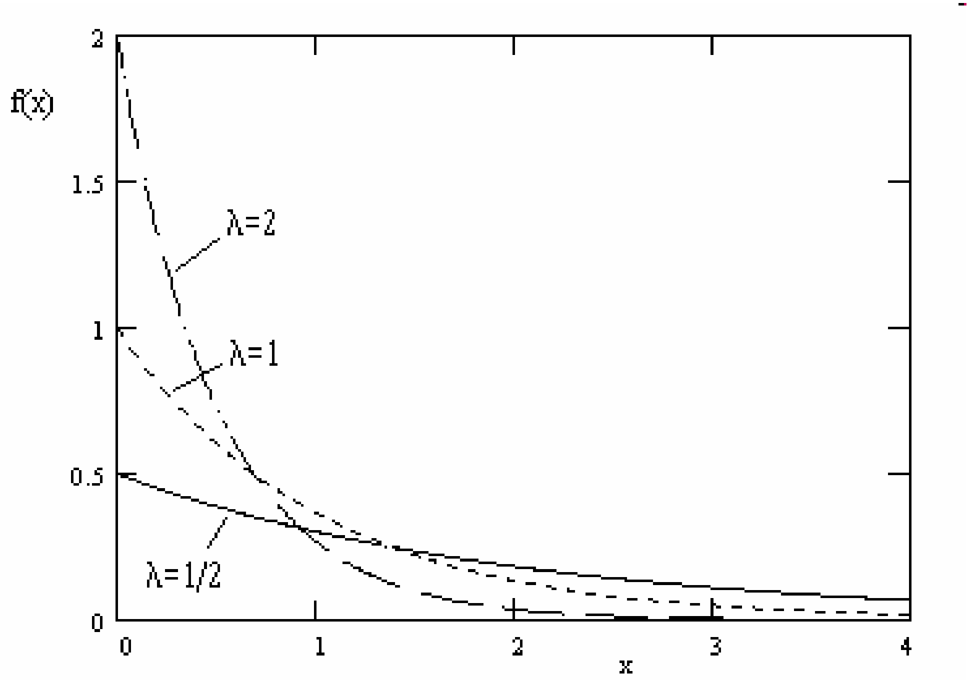


CONTENIDOS

1. Introducción a las colas poissonianas.
2. Modelo de colas poissoniano con un servidor *M/M/1*
3. Modelo con un servidor y capacidad finita *M/M/1/K*
4. Modelo con varios servidores *M/M/c*. Fórmula *C* de Erlang
5. Modelo con infinitos servidores *M/M/∞*
6. Modelo con varios servidores y pérdidas *M/M/c/K*
7. Modelo M/M/c con pérdidas (modelo *M/M/c/c*). Fórmula *B* de Erlang.
8. Modelo con un servidor, población finita *M/M/1/K/K*
9. Modelo con c servidores, población finita *M/M/c/K/K*

1. Introducción a los modelos de colas poissonianos

•Las colas poissonianas
(o exponenciales o markovianas)
son modelos del tipo *M/M*,
con llegadas de Poisson y
servicio exponencial,
que son las más estudiadas
analíticamente.



•Las llegadas de clientes y
su servicio demandado son
completamente aleatorios
en el sentido de que la
evolución del sistema depende
sólo de su estado
actual, y no de su pasado.

Introducción a las colas poissonianas

•Los procesos de nacimiento y muerte introducidos sirven para describir muchos modelos de colas.

•Asociaremos el término nacimiento con la llegada de un cliente al sistema y el término muerte con la salida de un cliente del sistema después de completado su servicio.

•El número de clientes en el sistema en el instante t, $N(t)$, indica el estado del mismo.

•Estudiaremos el comportamiento de las probabilidades $p_n(t)$ en el límite $\pi_n = \lim_{t\to\infty} p_n(t)$, que indica la proporción de tiempo que el sistema permanece con n clientes.

•La solución de equilibrio se obtenía de las ecuaciones que igualaban las tasas de entrada y salida de cada estado, dando lugar a

$$\pi_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \frac{\lambda_{n-1}\cdots\lambda_0}{\mu_n\cdots\mu_1}} \qquad \pi_n = \frac{\lambda_{n-1}\cdots\lambda_0}{\mu_n\cdots\mu_1}\pi_0, \quad n = 1, 2, 3\ldots \qquad \textbf{(10.1)}$$

•Para que exista dicha solución de equilibrio se debe satisfacer:

$$S_1 = 1 + \sum_{n=1}^{\infty} \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1} = \frac{1}{\pi_0} < \infty \qquad S_2 = \sum_{n=0}^{\infty} (1/(\lambda_n \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}})) = \infty$$

que ocurre si existe un n_0 tal que $\forall n > n_0$, $\lambda_n / \mu_n < 1$.

- Con los diversos λ_n, μ_n que se tendrán dependiendo del modelo en estudio:

  1. las ecuaciones de S_1 y S_2 servirán para buscar las condiciones bajo las que existe solución de equilibrio π_n, mientras que con

  2. las ecuaciones de π_0 y π_n obtendremos dicha solución.

Modelo: descripción, diagrama de transición, ecuaciones de equilibrio, análisis de la solución, medidas de rendimiento, tiempos de espera

2. Modelo de colas poissoniano con un servidor *M/M/1*

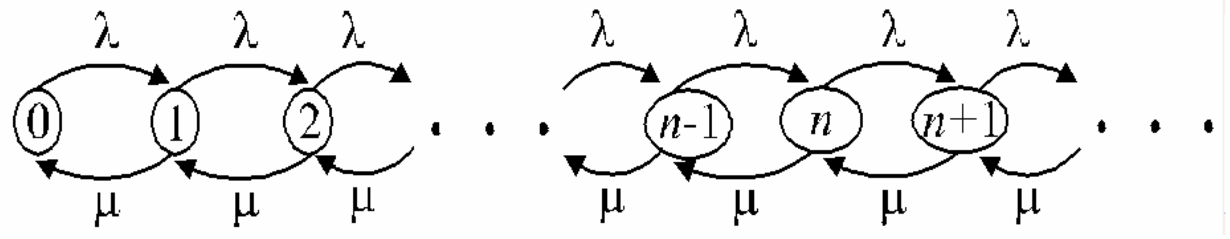
En este modelo se dispone sólo de un canal para dar servicio, las llegadas siguen un proceso de Poisson y la distribución del tiempo de servicio es exponencial.

Las tasas de nacimiento y muerte no dependen del número de clientes en el sistema y

$$\lambda_n = \lambda, \quad n = 0,1,2,\ldots \qquad \mu_n = \mu, \quad n = 1,2,3,\ldots$$

La capacidad del sistema es ilimitada y la disciplina de la cola es FIFO.

El diagrama de transición



Modelo de colas poissoniano con un servidor *M/M/1*

- El sistema de ecuaciones en equilibrio

Estado	Tasa entrada	=	Tasa salida
0	$\mu\pi_1$	$=$	$\lambda\pi_0$
$n \geq 1$	$\lambda\pi_{n-1} + \mu\pi_{n+1}$	$=$	$(\lambda + \mu)\pi_n$

$$
\begin{aligned}
\pi_0\lambda &= \pi_1\mu \\
\pi_1(\lambda+\mu) &= \pi_0\lambda + \pi_2\mu \\
\pi_2(\lambda+\mu) &= \pi_1\lambda + \pi_3\mu \\
&\cdots \\
\pi_i(\lambda+\mu) &= \pi_{i-1}\lambda + \pi_{i+1}\mu \\
&\cdots \\
\sum_{i=0}^{\infty}\pi_i &= 1
\end{aligned}
\quad\Rightarrow\quad
\begin{aligned}
\pi_0\lambda &= \pi_1\mu \\
\pi_1\lambda &= \pi_2\mu \\
\pi_2\lambda &= \pi_3\mu \\
&\cdots \\
\pi_i\lambda &= \pi_{i+1}\mu \\
&\cdots \\
\sum_{i=0}^{\infty}\pi_i &= 1
\end{aligned}
\quad\Rightarrow\quad
\begin{aligned}
\pi_1 &= \rho\pi_0 \\
\pi_2 &= \rho\pi_1 = \rho^2\pi_0 \\
\pi_3 &= \rho\pi_2 = \rho^3\pi_0 \\
&\cdots \\
\pi_i &= \rho\pi_{i-1} = \rho^i\pi_0 \\
&\cdots \\
\sum_{i=0}^{\infty}\pi_i &= 1
\end{aligned}
$$

Modelo de colas poissoniano con un servidor *M/M/1*

Sustituyendo las expresiones de los π_i en la última ecuación y despejando π_0 obtenemos (teniendo en cuenta que el factor de utilización es $\rho = r = \lambda/\mu$):

$$\pi_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n} = \frac{1}{1/(1-\rho)} = 1 - \rho$$

$$\pi_n = (1-\rho)\rho^n, \qquad n = 0, 1, 2, 3...$$

que corresponde a una distribución geométrica de parámetro 1- ρ.

$$S_1 = \sum_{n=0}^{\infty} \rho^n < \infty \quad S_2 = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{1}{\rho^n} = \infty$$

•La serie de S_1 converge si y sólo si $\rho < 1$.

•La segunda condición (S_2) se satis-face si $\rho \leq 1$.

•La condición necesaria y suficiente para que un modelo *M/M/1* tenga solución de equilibrio, es que $\rho<1$, que es la condición de estabilidad.

Modelo de colas poissoniano con un servidor *M/M/1*

•La probabilidad de que el canal esté ocupado es

$$P(\text{canal esté ocupado}) = 1 - \pi_0 = 1 - (1 - \rho) = \rho$$

•La probabilidad de encontrar al menos *n* de clientes en el sistema

$$P(N \geq n) = \sum_{k=n}^{\infty} \pi_k = \sum_{k=n}^{\infty} (1-\rho)\rho^k = (1-\rho)\rho^n \sum_{k=n}^{\infty} \rho^{k-n}$$
$$= (1-\rho)\rho^n \sum_{k=0}^{\infty} \rho^k = \frac{(1-\rho)\rho^n}{1-\rho} = \rho^n.$$

Modelo de colas poissoniano con un servidor *M/M/1*

Medidas de rendimiento

Comenzando por el número medio de clientes en el sistema, *L*, y en la cola, L_q. Se tiene

$$
\begin{aligned}
L = E(N) &= \sum_{n=0}^{\infty} n\pi_n = \sum_{n=0}^{\infty} n(1-\rho)\rho^n = (1-\rho)\sum_{n=0}^{\infty} n\rho^n \\
&= (1-\rho)\rho\sum_{n=1}^{\infty} n\rho^{n-1} = (1-\rho)\rho\frac{d}{d\rho}\left(\sum_{n=0}^{\infty}\rho^n\right) \\
&= (1-\rho)\rho\frac{d}{d\rho}\left(\frac{1}{1-\rho}\right) = \frac{(1-\rho)\rho}{(1-\rho)^2} = \frac{\rho}{1-\rho}.
\end{aligned}
$$

La sexta igualdad se debe a que las operaciones de suma y diferenciación pueden intercambiarse cuando las funciones implicadas se comportan lo suficientemente bien.

Otra expresión equivalente, en función de λ y μ, es $L = \dfrac{\lambda}{\mu - \lambda}$ ya que $\rho = \lambda/\mu$.

Modelo de colas poissoniano con un servidor *M/M/1*

L también podía haberse deducido directamente por tener *N* distribución geométrica.

Nótese que *L*, como función de ρ, tiene una asíntota vertical en $\rho=1$, lo que indica el dramático comportamiento del número medio de clientes en el sistema según nos acercamos hacia la violación de la condición de estabilidad.

Aparte de la media que acabamos de calcular de la variable *N*, podemos obtener su varianza, a partir de la distribución geométrica

$$\sigma_N^2 = \sum_{n=0}^{\infty} (n - L)^2 \pi_n = \frac{\rho}{(1-\rho)^2}$$

Calculamos el número medio de clientes en la cola L_q mediante

$$\begin{aligned} L_q = E(N_q) &= 0\pi_0 + \sum_{n=1}^{\infty} (n-1)\pi_n = \sum_{n=1}^{\infty} n\pi_n - \sum_{n=1}^{\infty} \pi_n \\ &= L - (1-\pi_0) = L - \rho = \frac{\rho^2}{1-\rho}, \end{aligned}$$

Modelo de colas poissoniano con un servidor *M/M/1*

que en términos de λ y μ es $$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Nótese que la igualdad $L_q = L - (1 - \pi_0)$ es general para cualquier cola con un servidor y dando servicio de uno en uno, ya que para obtenerla no se ha utilizado el tipo de distribuciones de los tiempos entre llegadas o de servicio.

Otra relación entre L y L_q es $$L_q = \frac{\rho}{1-\rho}\rho = L\rho = L\frac{\lambda}{\mu} \Longrightarrow \frac{L_q}{\lambda} = \frac{L}{\mu}$$

Recordemos que siempre $N = N_q + N_s$.
Pero en el modelo que estamos tratando, si $N \geq 1$, entonces $N = N_q + 1$, mientras que en general $L \neq L_q+1$, ya que L y L_q son medias y hay momentos en los que el servidor está desocupado.

Además, por las fórmulas de Little el número medio de clientes en el servidor $L_s = \lambda W_s = \rho$.

Calculamos el tamaño esperado de la cola cuando hay cola, denotado como $L_q' = E(N_q \mid N_q>0)$. Como la probabilidad condicionada de n clientes en el sistema dado que la cola no está vacía,

Modelo de colas poissoniano con un servidor *M/M/1*

En general, dadas dos v.a. X e Y, se verifica que

$$E(X) = E_Y[E(X \mid Y)].$$

Esto nos ofrece un camino alternativo para obtener L_q' a partir de L_q

$$\begin{aligned} L_q &= E(N_q) = E(N_q \mid N_q = 0)P(N_q = 0) + E(N_q \mid N_q > 0)P(N_q > 0) \\ &= E(N_q \mid N_q > 0)P(N_q > 0), \end{aligned}$$

$$E(N_q \mid N_q > 0) = \frac{L_q}{P(N_q > 0)} = \frac{L_q}{P(N \geq 2)} = \frac{L_q}{\rho^2} = \frac{1}{1-\rho}.$$

de donde

Modelo de colas poissoniano con un servidor *M/M/1*

$\pi_n' = P(n$ clientes en el sistema $\mid n \geq 2) =$

$P(n$ clientes en el sistema $, n \geq 2) / P(n \geq 2) =$

$\pi_n / (1- \pi_0 - \pi_1) =$

$\pi_n / \rho^2,$

para $n \geq 2$, se llega a

$$L_q' = E(N_q \mid N_q > 0) = \sum_{n=2}^{\infty}(n-1)\pi_n' = \sum_{n=2}^{\infty} n\frac{\pi_n}{\rho^2} - \sum_{n=2}^{\infty}\frac{\pi_n}{\rho^2}$$
$$= \frac{(L-\pi_1) - (1-\pi_0-\pi_1)}{\rho^2}$$
$$= \frac{(\rho/(1-\rho)) - (1-\pi_0)}{\rho^2} = \frac{1}{1-\rho}.$$

Tiempos de espera

Sólo resta estudiar los tiempos de espera del modelo *M/M/1*.
Obtendremos no sólo las medias W y W_q,
sino también las distribuciones de probabilidad de las v.a. w y q.

Las medias se calculan fácilmente por las fórmulas de Little:

$$W = E(w) = \frac{L}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{E(s)}{1-\rho}$$

$$W_q = E(q) = \frac{L_q}{\lambda} = \frac{\rho^2}{\lambda(1-\rho)} = \frac{\rho E(s)}{1-\rho}$$

Nótese que como dijimos para L, W tiene también un comportamiento dramático según ρ tiende a 1.

Ahora queremos calcular el tiempo medio de espera en cola para aquellos clientes que deben esperar.

Modelo de colas poissoniano con un servidor *M/M/1*

Como

$$\begin{aligned} W_q &= E(q) = E(q \mid q = 0)P(q = 0) + E(q \mid q > 0)P(q > 0) \\ &= 0(1 - \rho) + E(q \mid q > 0)\rho, \end{aligned}$$

se tiene

$$E(q \mid q > 0) = \frac{W_q}{\rho} = \frac{E(s)}{1 - \rho} = W$$

Esta cantidad interesa porque un tiempo medio de espera aceptable puede deberse a que muchos clientes no tienen que esperar, pero los que esperan lo hacen durante mucho tiempo.

Como $W = W_q + E(s)$, tenemos que $E(q \mid q > 0) = W_q + E(s)$, lo que indica que, en media, los clientes que tienen que esperar en cola esperan más que lo que un cliente medio espera, ya que espera un tiempo medio de servicio, $E(s)$, más.

Modelo de colas poissoniano con un servidor *M/M/1*

Para hallar la distribución de la variable aleatoria q, nótese primero que tiene un punto (t = 0) con probabilidad positiva: $P(q = 0) = P(N = 0) = 1 - \rho$.

Por otra parte, si al llegar el cliente encuentra n personas en el sistema (una de ellas en el servidor), tendrá que esperar a que todas se sirvan.
Así, el tiempo de espera en cola es la suma de las variables aleatorias "tiempo de servicio del cliente i ", $i = 1,\dots,n$,

$$q = s_1 + \ldots + s_n$$

en donde s_i son independientes e idénticamente distribuidas según una exponencial de parámetro μ.

Debemos recordar que por la pérdida de memoria de la distribución exponencial, no hace falta tener en cuenta el tiempo de servicio ya consumido por el cliente que actualmente está sirviéndose.

Por la reproductividad de la distribución gamma, $q \mid N=n$ sigue una distribución gamma de parámetros $p = n$, a = μ (en este caso es Erlang, al ser p entero).

Modelo de colas poissoniano con un servidor *M/M/1*

Por el teorema de la probabilidad total se tiene

$$\begin{aligned}
P(0 < q \le t) &= \sum_{n=1}^{\infty} P(q \le t \mid N = n) P(N = n) \\
&= \sum_{n=1}^{\infty} \int_0^t \mu e^{-\mu x} \frac{(\mu x)^{n-1}}{(n-1)!} \rho^n (1-\rho) dx \\
&= \int_0^t (1-\rho) e^{-\mu x} \sum_{n=1}^{\infty} \frac{(\mu \rho x)^{n-1}}{(n-1)!} \rho \mu dx \\
&= \int_0^t \mu \rho (1-\rho) e^{-\mu x} e^{\mu \rho x} dx = \rho(1 - e^{-\mu(1-\rho)t}) = \rho(1 - e^{-t/W}).
\end{aligned}$$

Luego, la función de distribución de *q* es

$$F_q(t) = P(q = 0) + P(0 < q \le t) = 1 - \rho + \rho(1 - e^{-t/W}) = 1 - \rho e^{-t/W}.$$

Esta expresión es válida para todo *t*, aunque *q* sea discreta en el origen y continua para *t* > 0.

Modelo de colas poissoniano con un servidor *M/M/1*

Análogamente, calculamos la distribución de la v.a. *w*.
Si cuando llega un cliente ya hay *n* en el sistema, éste tendrá que estar en el sistema un tiem-po total igual a la suma de *n* + 1 v.a.i.i.d. según una ley exponencial de parámetro μ.

Así, la distribución de *w* será una gamma de parámetros *p* = *n*+1, *a* = μ.
Variando *n*, por el teorema de la probabilidad total:

$$
\begin{aligned}
F_w(t) &= P(w \le t) = \sum_{n=0}^{\infty} P(w \le t \mid N = n)P(N = n) \\
&= \sum_{n=0}^{\infty} \int_0^t \mu e^{-\mu x} \frac{(\mu x)^n}{n!} \rho^n (1-\rho) dx = \int_0^t \mu(1-\rho) e^{-\mu x} \sum_{n=0}^{\infty} \frac{(\mu \rho x)^n}{n!} dx \\
&= \int_0^t \mu(1-\rho) e^{-\mu x} e^{\mu \rho x} dx = 1 - e^{-\mu(1-\rho)t} = 1 - e^{-t/W} \quad (t \ge 0).
\end{aligned}
$$

Es decir, *w* sigue una distribución exponencial de parámetro
μ(1 – *ρ*) = μ –λ = 1/W.